



Spatio-Temporal Tensor Decomposition of a Polyaffine Motion Model for a Better Analysis of Pathological Left Ventricular Dynamics

Kristin Mcleod, Maxime Sermesant, Philipp Beerbaum, Xavier Pennec

► To cite this version:

Kristin Mcleod, Maxime Sermesant, Philipp Beerbaum, Xavier Pennec. Spatio-Temporal Tensor Decomposition of a Polyaffine Motion Model for a Better Analysis of Pathological Left Ventricular Dynamics. IEEE Transactions on Medical Imaging, 2015, 34 (7), pp.1562–1675. 10.1109/TMI.2015.2405579 . hal-01205342

HAL Id: hal-01205342

<https://inria.hal.science/hal-01205342>

Submitted on 25 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spatio-Temporal Tensor Decomposition of a Polyaffine Motion Model for a Better Analysis of Pathological Left Ventricular Dynamics

Kristin McLeod, Maxime Sermesant, Philipp Beerbaum, Xavier Pennec

Abstract—Given that heart disease can cause abnormal motion dynamics over the cardiac cycle, which can then affect cardiac function, understanding and quantifying cardiac motion can provide insight for clinicians to aid in diagnosis, therapy planning, as well as to determine the prognosis for a given patient. The goal of this paper is to extract population-specific cardiac motion patterns from 3D displacements in order to firstly identify the mean motion behaviour in a population and secondly to describe pathology-specific motion patterns in terms of the spatial and temporal aspects of the motion. Since there are common motion patterns observed in patients suffering from the same condition, extracting these patterns can lead towards a better understanding of a disease. Quantifying cardiac motion at a population level is not a simple task since images can vary widely in terms of image quality, size, resolution and pose. To overcome this, we analyse the parameters obtained from a cardiac-specific Polyaffine motion tracking algorithm, which are aligned both spatially and temporally to a common reference space. Once all parameters are aligned, different subjects and different populations can be compared and analysed in the space of Polyaffine transformations by projecting the transformations to a reduced-order subspace in which dominant motion patterns in each population can be extracted and analysed. Using tensor decomposition allows the spatial and temporal aspects to be decoupled in order to study the different components individually. The proposed method was validated on healthy volunteers and Tetralogy of Fallot patients according to known spatial and temporal behaviour for each population. A key advantage of the proposed method is the ability to regenerate motion sequences from the respective models, thus the models can be visualised in terms of the full motion, which allows for better understanding of the motion dynamics of different populations.

Index Terms—cardiac modelling, atlas, non-rigid image registration, motion tracking, population statistics, tensor decomposition, spatio-temporal alignment, Tetralogy of Fallot.

I. INTRODUCTION

Cardiovascular disease is a worldwide issue, being the leading cause of death and affecting the day-to-day life of millions of individuals. Heart disease can affect the motion

dynamics of the heart over the cardiac cycle. Understanding the abnormal dynamics can thus potentially give insight into a pathology and provide information that can aid with diagnosis, therapy planning, and for determining the prognosis for a given patient. Quantitatively defining what are “normal” motion patterns is not straightforward given the complex dynamics and coupling between the ventricles and atria. In patients with heart conditions that affect the shape of the ventricles, the motion dynamics can be affected by the abnormal shape, resulting in poor pumping function in one or both of the ventricles, ultimately leading to increased workload on the heart. For patients with stenosis in the arteries or poor valve function, the motion dynamics can be affected by the excess load of the heart trying to pump blood to get enough oxygen to the body. Given the heavy workload on the heart in a lifetime, maintaining healthy heart function is crucial to maximise longevity.

The motion of the heart can be measured by placing markers on the myocardial tissue, which are then tracked over time. However, this kind of measure is invasive and not applicable on a large-scale basis or in all populations. For this reason, medical images are commonly used to non-invasively visualise the motion of the heart as it beats using magnetic resonance imaging (MRI), echocardiography (ECHO), or computed tomography (CT), for example. Using such images, the temporal evolution of the heart can be visualised in 2D or 3D. Such image sequences can be used to qualitatively analyse the motion dynamics. Some medical scanners are also able to compute global measures directly from the images such as ventricular volume, or 1D blood flow measures, however, 3D measures are not currently available in standard scanners. Also, more in-depth measures of cardiac function cannot be extracted directly from the images and quantitatively comparing motion in different subjects beyond strain or global measures is an ongoing challenge.

Quantitatively tracking cardiac motion can provide further insight into the 3D dynamics to obtain measures of cardiac function by analysing how efficiently the heart is beating and by identifying patterns related to normal (healthy) motion, and patterns characteristic of a given pathology. Myocardial deformation and strain can give early indicators of poor function and disease evolution. A number of methods for tracking cardiac tissue in a single subject have been proposed using image registration techniques to compute the deformation between images in the cardiac cycle. These methods are further discussed in Sec. I-1. Methods for population-based analysis

Copyright ©2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

K. McLeod is with the Cardiac Modelling department of Simula Research Laboratory, Norway, the Centre for Cardiological Innovation (CCI), Norway and the Asclepios Research Group, INRIA, Sophia Antipolis, France (e-mail: kristin@simula.no)

P. Beerbaum is with Hanover Medical University, Germany (e-mail: Beerbaum.Phillipp@mh-hannover.de)

M. Sermesant and X. Pennec are with the Asclepios Research Group, INRIA, Sophia Antipolis, France (e-mail: maxime.sermesant@inria.fr, xavier.pennec@inria.fr)

Manuscript received May xx, 2014; revised December xx, 2014.

of cardiac motion have been proposed to study the functional dynamics of the heart at a population-wide level. A brief introduction of these methods is given in Sec. I-2.

1) *Cardiac Motion Tracking using Non-rigid Registration:* A number of non-rigid image registration techniques for cardiac motion tracking have been proposed in recent years. A recent review of these is given in [1], [2]. Key features of these methods include ensuring that the deformations are diffeomorphic (smooth transformations that preserve the structure of the material to prevent non-physiological transformations such as folding) and including an incompressibility term to prevent large volume change in the myocardium. The challenge in cardiac motion tracking of cine MRI images is that the tissue is represented homogeneously in the images, giving little texture information to track besides the endocardial and epicardial contours. Thus, it is difficult to capture the strain purely from image-driven registration algorithms, particularly the circumferential strain. Voxel-based methods, which essentially model the correspondences between voxels in pairs of images, have been widely used for cardiac registration [3],[4]. These methods have the advantage of defining the motion in every voxel in the image. However, since the displacement fields are dense, the number of parameters needed to describe the motion is large. For comparative population-wise studies, such a parameterisation reduces the robustness in group-wise analysis since comparing large parameter sets is difficult. Methods parameterised by B-spline transformations on the other hand have a lower number of parameters since the deformation is parameterised by the motion of control points, hence such methods have been used for cardiac motion tracking, for example in [5], [6], [7], [8]. However, from B-spline based tracking, the total number of parameters relies on the sampling of control points, which is a trade-off between the accuracy of the tracking and maintaining a low number of parameters, since the motion is defined only at the control points. Simple models that can accurately represent the full deformation with a small number of parameters could provide a more insightful basis for motion analysis. The Polyaffine model [9], [10] provides such a framework by combining regional affine transformations in a smooth, global manner and has previously been applied in cardiac and cardiovascular image registration [11], [12], [13], [14]. The Polyaffine model has the advantage of providing a regional-level analysis of the motion, and since the deformation parameters are affine, the parameters themselves are consistently defined from one sequence to another.

2) *Population-based Motion Analysis:* In [15], myocardial strain extracted from tagged MRI was used to identify and localise regional abnormal cardiac motion patterns, by describing the spatio-temporal motion as a tensor, where the tensor contains regional strain information for 20 regions of the left ventricle. This results in a tensor of size 60×10 (3 strain directions \times 20 regions \times 10 frames from end diastole to end systole) that is used to classify healthy/abnormal motion patterns. A key advantage of this work is the analysis of the parameters separately for space and time using 2D linear discriminant analysis (2DLDA) and a k-nearest-neighbour method to classify subjects as healthy or diseased.

The method provides an interesting framework for population-specific function analysis, but is limited to 1D strain analysis. Therefore, some of the more complex 3D dynamics are not described with this model. Displacement fields were compared across subjects in [16] by first extracting the motion parameters using a B-spline model, then transforming the motion field to a common space by transforming each displacement vector by the inverse Jacobian matrix of the transformation. This was extended to construct a cardiac motion atlas in [17] by simply averaging the transformed motion fields from a group of subjects drawn from the same population. A 4D statistical atlas of the cardiac anatomy was developed using a B-spline model to represent the transformations in [18] by first performing spatio-temporal alignment [19]. This model was used to differentiate between normal motion and the motion of patients with hypertrophic cardiomyopathy. A cardiac motion atlas was constructed in [20] by performing principal component analysis (PCA) on subject-specific motion fields after first spatio-temporally aligning the fields. Regional measures are then computed by averaging over the 17 American Heart Association (AHA) regions of the left ventricle.

A. Aim and Paper Organisation

Previous methods for performing population-based motion analysis have been directed towards analysis of dense motion fields, or of regional strain values. In terms of comparing 3D displacements, the analysis may require alignment of either images or displacement fields both spatially and temporally. In general, the spatio-temporal alignment of either images or displacement fields creates a bottleneck in population-based analysis due to the difficulty of aligning high dimensional objects. A key point missing in previous population-based work is further analysis towards understanding motion abnormalities (beyond identifying or classifying motion abnormalities).

In the present work, we are interested in performing group-wise analysis of a small set of displacement parameters obtained from the cardiac-specific PolyLogDemons algorithm of [13], [14] as well as including image-driven spatial features to the model. We are particularly interested in analysing motion from short axis cine-MR images, given that this image sequence is commonly used in routine clinical practice (and thus cine-MR images are more readily available), and has the advantage of capturing the motion with reasonable spatial resolution (compared to ECHO for example). Rather than performing PCA on motion fields as in [21], or tensor decomposition of regional strain values as in [15], we perform analysis on Polyaffine parameters, and compute spatial and temporal bases by performing Tucker tensor decomposition on the parameters of a set of subjects. In contrast to the method proposed in [22] that was proposed to represent high-dimensional deformation fields in the context of brain image warping by a 3D statistical representation, we represent the deformation fields directly with Polyaffine transformation parameters.

Preliminary versions of this work have been published describing the motion tracking [13], analysis of the trace of affine matrices [14], and a preliminary investigation of matrix-based parameter decomposition to analyse spatial and temporal

behaviour [23]. The main contributions of the present work extending on these previous articles are summarised below:

- Image-driven spatial priors were added to [13], [14]
- A framework for spatio-temporal alignment of Polyaffine parameters is proposed
- A method for the construction of a generative mean motion model is proposed
- Tensor decomposition of Polyaffine parameters is proposed to compare dominant spatial and temporal features between populations

The remainder of this paper is organised as follows. The methods section (Sec. II) begins with a brief introduction of the cardiac-specific PolyLogDemons algorithm of [14] that is used to obtain the transformation parameters that are used in the population-based analyses (Sec. II-A). The additional image-based model priors that are added to the model are described in Sec. II-A1. The proposed spatio-temporal alignment of all Polyaffine parameters is described in Sec. II-B. Methods for performing population-based motion analysis using Tensor decomposition of Polyaffine parameters are proposed in Sec. II-C. The validation section (Sec. III) describes first; the validation of the motion tracking (registration) using ground-truth motion landmarks (Sec. III-A), second; the validation of the mean motion model in Sec. III-B, third; an analysis of the predictive power of the mean models (Sec. III-C) and finally; validation of the tensor-based analysis (Sec. III-D).

II. METHODS

A. Transformation Parameter Estimation using a Polyaffine Model

For a given set of regions, the transformation in each region i can be modelled by an affine deformation parameterised by a 3×4 matrix M_i . The M_i matrices can be fused to a global deformation field using the Polyaffine model:

$$\vec{v}_{poly}(x) = \sum_i \omega_i(x) M_i \tilde{x}, \quad (1)$$

where ω_i is a parameter controlling the weight of the i^{th} region for each voxel x written in homogeneous coordinates [24].

As shown in [25], Eq. 1 can be estimated by a linear least squares projection from an observed velocity field $\vec{v}(x)$ (in this case computed using the LogDemons algorithm) to the space of Log-Euclidean Polyaffine Transformations (LEPT's). The log affine parameters M_i can be estimated by the following least-squares regression:

$$C(M) = \int_{\Pi} \left\| \sum_i \omega_i(x) \cdot M_i \tilde{x} - \vec{v}(x) \right\|^2 dx.$$

Using the Frobenius inner product $\|W\|^2 = Tr(W^T W)$, the solution at the optimum $\nabla C_M = 0$ is given by $M = B \cdot \Sigma^{-1}$ [25], which in vector form is equivalently $vect(M) = (\Sigma \otimes Id_3)^{-1} \cdot vect(B)$, where $M = [M_1 M_2 \dots M_3]$, $B_i = \int_{\Pi} \omega_i(x) \cdot \vec{v}(x) \cdot \tilde{x}^T dx$ and $\Sigma_{ij} = \int_{\Pi} \omega_i(x) \cdot \omega_j(x) \cdot x \cdot x^T dx$.

In [13], the authors showed that an incompressibility penalisation and a neighbouring region regularisation can be added to this model by penalising the trace of the gradient of the Polyaffine velocity field, and by regularising the similarity

between neighbouring regions via a distance term. Both terms were added in the least squares minimisation to obtain a penalised least squares error term, still with a linear solution in terms of M . The penalisation term was derived as:

$$\alpha \int_{\Pi} Tr(\nabla \vec{v}_{poly})^2 dx, \quad (2)$$

where the parameter α is used to control the strength of the penalisation. An additional term was added for the regularisation:

$$\beta \sum_i \Omega_{ij} dist(M_i, M_j), \quad (3)$$

where β controls the strength of the regularisation. Ω_{ij} is defined as:

$$\Omega_{ij} = \int_{\Pi} \omega_i(x) \omega_j(x) dx. \quad (4)$$

The new solution for M is given by:

$$vect(M) = (\Sigma \otimes Id_3 + \alpha V + \beta R)^{-1} \cdot vect(B), \quad (5)$$

where R and V are the matrices controlling the regularisation and incompressibility respectively, as described in [13].

The Polyaffine regions used in [13], [14] were defined as cardiac-specific regions for the left ventricle using the 17 American Heart Association (AHA) regions. In [14], the Polyaffine weight functions were computed in prolate spheroidal (PSS) coordinates rather than in the Cartesian frame using the method of Toussaint et. al [26] (see Fig. 1). Using the prolate spheroidal coordinates provides more anatomically grounded shapes of the weights to smooth the Polyaffine fusion around the ventricle.

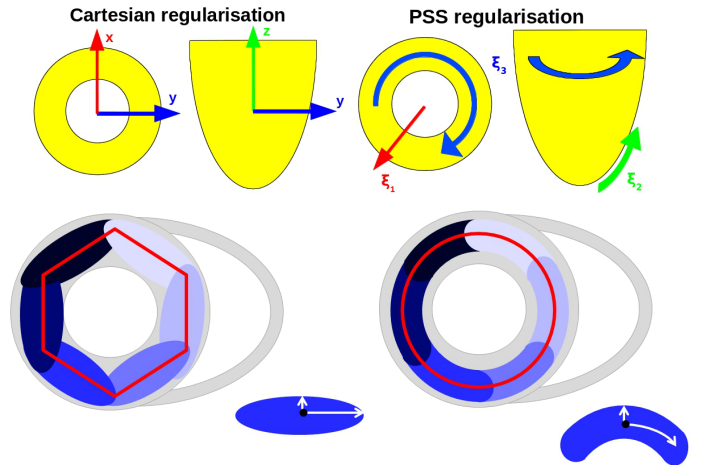


Fig. 1. Regularisation direction for Cartesian coordinates (top left) and prolate spheroidal (PSS) coordinates (top right). The difference between Cartesian and PSS weights shown from above on a spheroid and the directions of the covariance matrix shown on an oval (bottom left - lower right corner) and curved segment (bottom right - lower right corner). The red lines emphasise the different shapes captured by the different weights, where the Cartesian weights sum to a hexagon (left) and the PSS weights sum to a circle (right).

1) Image-based Priors:

a) *Image Gradient Based Velocity Weighting*: Image-based spatial priors can be included in the model to weight the input velocity field to give more influence on velocity vectors in regions of high gradient in the image (regions that drive the LogDemons registration). In [27], the region of interest, Π , was defined as a binary mask, where $\Pi = 1$ inside the mask, and $\Pi = 0$ outside. In this way, equal weight is given to the inner part of the tissue in the myocardium. In cine MR images, however, there is little texture information in the inner part of the tissue. Therefore, the boundary of the tissue is the region that drives the registration, since this is the region of highest gradient. In fact, given the Gaussian nature of the Polyaffine weights, more weight is given to the centre of the tissue than the boundaries. To account for this, we introduce a spatial prior on the LogDemons correspondence field by giving higher confidence to the vectors in regions with high gradients in the image. This prior can be easily incorporated into the model by modifying the mask over which we integrate the terms in Eq. II-A (Π). Rather than using a standard binary mask, we can incorporate a confidence on the input velocity field (the LogDemons correspondence field) by taking the smoothed gradient magnitude image of the fixed and moving images (or similarly, we could use the structure tensor as used in [28]). For the reference image, however, the gradient of the binary mask image can be used since this is already defined (since it is needed for the computation of the weights). By taking the addition of these two images for the fixed and moving frames, we obtain an image with the added confidence between the fixed and moving images, which can be used to weight the input velocity field to give higher confidence to the velocities in voxels on the boundary of the tissue.

b) *Epicardial constraints*: A second prior can be added to restrict the displacement in the epicardium, preventing the registration from being dragged by the endocardial motion. This is to account for the fact that the epicardial borders remain more fixed due to the constraint on the motion from the pericardium. This prior can be included in the model by down-weighting the input velocity field at the epicardium by a factor τ , where $0 < \tau < 1$, and $\tau = 1$ in regions of high gradient (thus leaving the velocity vectors unchanged). We can define τ as $(1 - (1 - C(x)) \times E(x))$, where $E(x)$ defines the epicardial border, and $C(x)$ is the magnitude of the image gradient. In practice, this can be the gradient magnitude image of the epicardial mask. The term $1 - C(x)$ is included to prevent the velocity from being restricted in regions where there is high gradient in the image (and thus in regions where we should trust the velocity).

Using this cardiac-specific PolyLogDemons method, a low-dimensional, regional, and consistent definition of the motion can be obtained for a different subjects. Moreover, the obtained deformations are physiologically grounded and the method makes use of image-specific intensity features.

B. Spatio-Temporal Parameter Alignment

Given the wide variability of cine MRI data, including temporal shifts, different cycle lengths, and different coordinate spaces for different acquisitions, spatio-temporal alignment

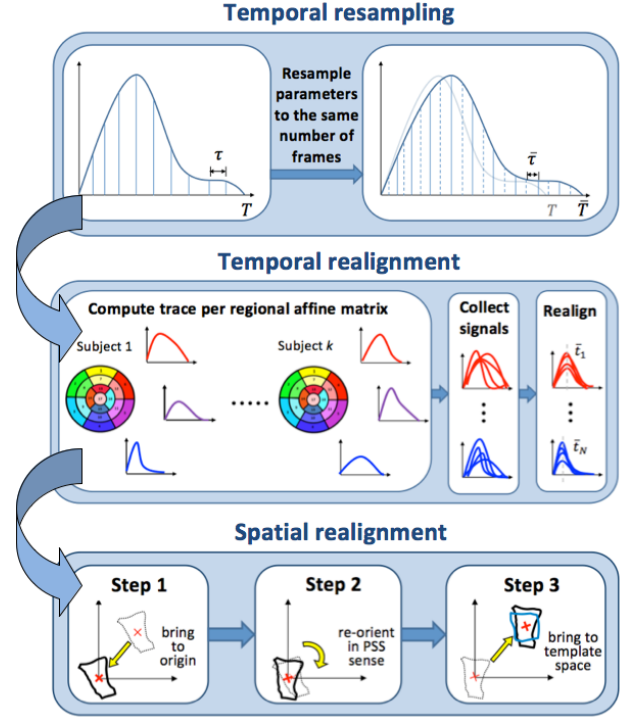


Fig. 2. Spatio-temporal alignment pipeline starting with a temporal resampling step to resample the parameters so that there are an equal number of frames for all subjects. Temporal realignment is then performed regionally to align the parameters by the regional mean peak of the trace of the affine matrix. Spatial alignment is performed to align all parameters to a common coordinate space, by transporting the parameters to be centred at the barycentre of the region for the reference, and reoriented in a prolate spheroidal sense.

is needed. We propose to perform the alignment directly on the transformation parameters to align all the parameters to a common space rather than pre-aligning the images or post-aligning the displacements. A method for temporally resampling the parameters so that all subjects have the same number of frames is proposed in Sec. II-B1, followed by a method for temporally aligning the parameters in Sec. II-B2, and finally a method for spatially aligning the parameters to a common space in a prolate spheroidal sense in Sec. II-B3. The steps for aligning Polyaffine transformation parameters are summarised in Fig. 2.

1) *Temporal Resampling*: In cine MR imaging, there is not a fixed number of frames acquired for each subject. In practice, the cycle length (from the R peak to the next R peak) is divided into a given number of frames (typically 10-30), and the cine movie is constructed from the information gathered over several heart beats (see Fig. 3). Such sequences are acquired while subjects hold their breathe, to remove lung motion artifacts. Therefore, since the cycle length differs from one acquisition to another, we need to temporally resample the parameters to a common space to have an equal number of frames for all subjects (in order to be able to directly compare the parameters). This is done simply by interpolating between frames to resample the number of observations to be the same for all subjects. This is represented in Fig. 2, step 1.

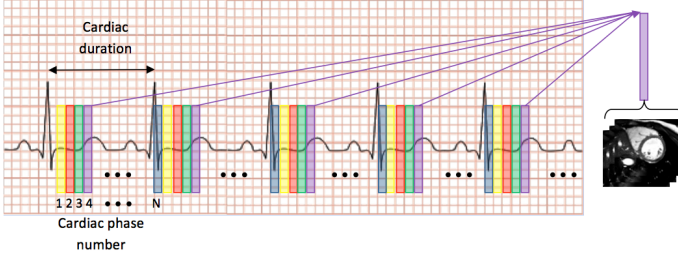


Fig. 3. For a cine image sequence, the cardiac cycle is divided into N phases and the acquisition is averaged for each phase over several cycles. The number of phases is variable from one acquisition to another.

2) *Regional Temporal Alignment:* Cine MR image sequences are generally gated by electrocardiogram (ECG) signals, so that the first frame corresponds to the end diastolic frame, and so that an image sequence covers one cardiac cycle. Standard acquisition is acquired by waiting a given amount of time after the R wave before initiating the sequence (see Fig. 4) and images are acquired between the R -to- R cycle. However, sometimes the first frame does not correspond exactly to the end diastolic frame and thus the sequence is slightly offset. Furthermore, some sequences do not loop back exactly to the end diastolic frame, where some frames may be missing, or too many frames may be included in the sequence. Also, the ECG signal may be weak in some patients, such as those with large pericardial effusion, making the ECG gating difficult. Subjects may also differ in the duration of the cardiac cycle, therefore, the peak frame does not always overlap for a group of patients. To account for this, the transformation parameters can be aligned temporally per region, by computing the trace of the affine matrix per region over time. The trace of the affine matrix corresponds to the volume curve, through properties of Lie algebra, so this essentially amounts to aligning the parameters according to the volume curves. In this way, the mean peak frame can be computed for a group of patients, and all the transformation parameters can be aligned to the mean peak frame (see Fig. 5). We assume that the transformation starts and finishes at the identity transformation. A temporal shift in one patient of five frames would set the first five frames to the identity transformation. This has the effect of not implying any information for these frames when the values are unknown, and thus the mean for these frames is computed only from the subjects in which the transformation is given.

3) *Spatial Alignment:* In order to be able to construct a mean motion model, we first need to transport all the parameters to a common space. Rather than aligning all the images, which would amount to rigidly registering all images to a common space prior to performing the motion tracking, as was done in [23], we rather align the transformations after performing the motion tracking. Given the low dimensional representation of the motion given from the registration algorithm summarised in the previous section, this is a matter of simply transporting the affine matrices to a common coordinate space. By choosing one subject (at random) to be the reference subject, we can transport all the other subjects to this space. We are interested in analysing

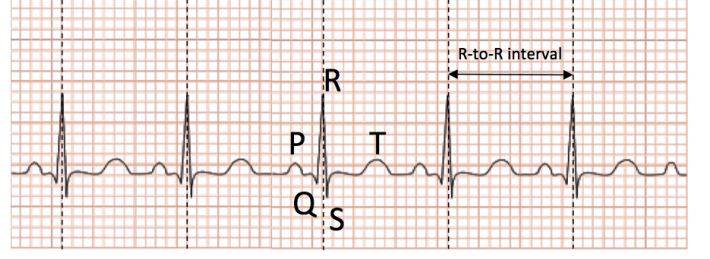


Fig. 4. The ECG signal is used to detect the R peak. Acquisition is started a fixed amount of time after the R peak, and one cycle is considered as the R -to- R duration. Gating the cine MR sequences in this way can result in different acquisition lengths, especially for patients with low ECG signal.

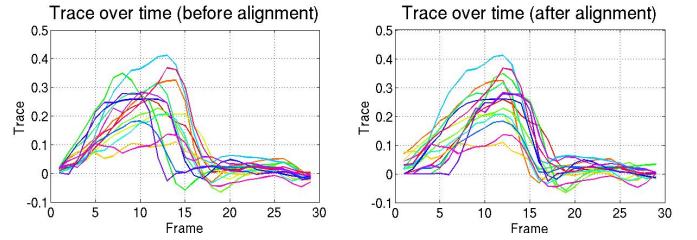


Fig. 5. Trace of the affine matrix for a single region plotted over time for a group of subjects shown before alignment (left) and after alignment (right). Linearly aligning the curves brings the peaks to the same frame, to give the amount of shift needed to align the parameters.

the motion at a regional level, hence we are interested in aligning the matrices at a regional level independently from other regions. Therefore, we can align all matrices by first re-centring each matrix around the origin using the barycentre per region. The regions then need to be re-oriented to be aligned in a common manner. Since the Polyaffine weights are defined in PSS coordinates, it follows that the matrices can be oriented in a prolate spheroidal sense (see Fig. 1) by taking the Jacobian matrix computed at the barycentre. The same approach is applied to re-orient the result to match the template orientation, followed by a translation to bring the matrix to the template space. A simplification of this pipeline is shown in Fig. 2, step 3. The prolate spheroidal coordinates were computed using the method of Toussaint et. al [26].

C. Population-based Motion Analysis

1) *Mean Motion Model Construction:* Once all the matrices are aligned to a common space for all subjects, the mean of these observations can be easily computed using basic arithmetic operations. The mean \bar{M}_t at time t can be computed by:

$$\bar{M}_t = \frac{1}{N_k} \sum_{i=1}^{N_k} M_{k,t}, \quad (6)$$

where k is the subject index, N_k is the number of subjects, and $M_{k,t}$ is the transported matrix for subject k at time t . By stacking all the matrices of a given subject to form a column vector of [time \times number of regions \times affine parameters], we can similarly compute the mean trajectory \bar{M} by taking the average of these vectors.

2) *Simulated Motion from the Mean Model*: Given the estimated mean trajectory, the motion can be simulated on a new subject, given that the subject is in the same coordinate space as the reference, by simply applying the mean \bar{M}_t to the first image of the sequence: $I_t = \bar{M}_t * I_0$ (see Fig. 6).

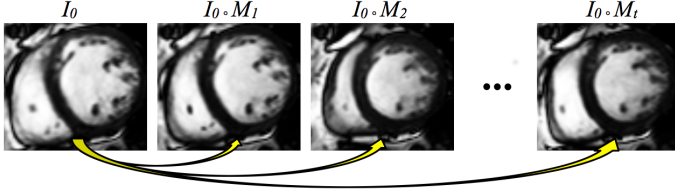


Fig. 6. The mean motion model applied to an image requires simply applying the mean transformation per frame to the reference end-diastolic image.

In general, new subjects will not be in the same coordinate space as the template. Therefore, the mean trajectory \bar{M}_t first needs to be transported to the subject space. This can be done in a similar way to the pipeline for aligning the subjects. Given the barycentre and Jacobian matrix for each region of the new subject, \bar{M}_t can be transported to the subject space by:

$$M_r^k = T_r^k \cdot (\bar{T}_r)^{-1} \cdot \bar{M}_r \cdot \bar{T}_r \cdot (T_r^k)^{-1}, \quad (7)$$

where M_r^k is the mean motion model in region r in the subject-specific space for subject k , T_r^k is the matrix:

$$T_r^k = \begin{bmatrix} J_r^k & b_r^k \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix}$$

with b_r^k the barycentre of region r for the k^{th} subject (in Cartesian coordinates), and J_r^k is the Jacobian matrix at the barycentre in the local (prolate spheroidal) coordinates. \bar{T}_r is equivalently defined for the mean at the barycentre of region r in the reference space.

D. Spatial and Temporal Mode Construction using Tensor Decomposition

Rather than performing PCA on stacked matrices of all parameters for each subject or singular value decomposition (SVD) on parameters stacked either temporally or spatially as in [23], the tensor \mathcal{T} of all parameters can be decomposed directly. A number of methods for tensor decomposition have been proposed, the most common being the CANDECOMP/PARAFAC (CP) decomposition [29], [30] and the Tucker decomposition [31], [32] (see Fig. 7). A review of these methods, as well as other tensor decomposition methods, is given in [33]. The CP decomposition is the generalisation of SVD to tensors, in which a tensor is factored into the sum of rank-one tensors (tensors that can be written as the outer product of vectors). The Tucker decomposition (also known as higher-order PCA), rather decomposes the tensor into a core tensor and factor matrices corresponding to each axis (see Fig. 7). In this work, we focus on the Tucker decomposition, to give decoupled “loadings” of the modes (whereas in the case of the CP decomposition, there is a single loading for each axis, meaning that the n^{th} modes are coupled in some way).

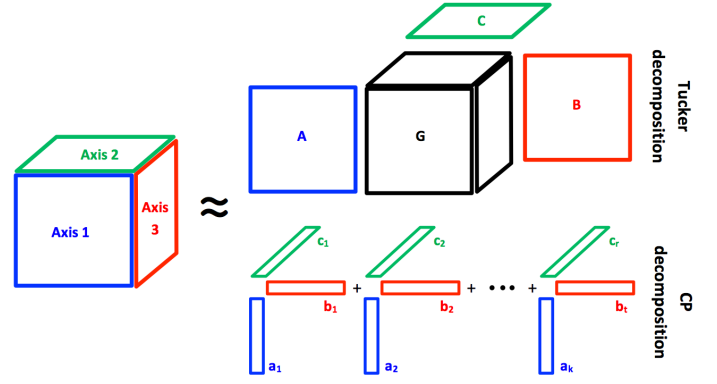


Fig. 7. Visual representation of the Tucker decomposition (top right) and CP decomposition (bottom right) on a three-way tensor.

1) *Tensor Decomposition Using the Tucker Method*: The Tucker decomposition (a.k.a higher-order PCA/SVD, N-mode SVD/PCA) computes an orthonormal space associated with each axis of the tensor \mathcal{T} . For a three-way tensor, the Tucker decomposition of \mathcal{T} is expressed as an n-mode product:

$$\begin{aligned} \mathcal{T} \approx \mathcal{G} \times_1 A \times_2 B \times_3 C &= \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} a_p \circ b_q \circ c_r \\ &= [[\mathcal{G}; A, B, C]], \end{aligned} \quad (8)$$

where \times_n denotes the mode- n tensor-matrix product, and \circ denotes the vector outer product. A , B , and C are loading matrices in each direction that can be thought of as the tensor equivalent of principal components for each axis. The core tensor \mathcal{G} gives the relationship between the modes in each direction and describes how to re-combine the original tensor \mathcal{T} . Computing such a decomposition is not straightforward. One method for doing so is to matricise the tensor (in the three-way case), similar to the decomposition described in [23]. Another method, known as the *Tucker1* method, is to compute optimal components to capture the variation in the n^{th} axis, independently from the other axes. Alternatively, the factors can be estimated using an iterative alternating least squares strategy (ALS), an alternating slice-wise decomposition (ASD), non-linear conjugate gradient method (NCG) and using a direct multi-linear decomposition (DMLD). Further details on the ALS strategy used in this work for computing a Tucker decomposition is given in the Appendix.

The tensor decomposition can be performed on a tensor stacked in a number of different ways, and would provide different levels of analysis of the parameters. In order to compute the spatial and temporal modes separately, all the spatial parameters can be stacked as a single vector of all affine parameters for all regions, along time, and for all subjects (namely a 3-way tensor of *space* \times *time* \times *subject*). This decomposition enables separate spatial and temporal mode analysis to be performed so that the reduced-order model can be validated by comparing to known spatial and temporal dynamics. It is also possible to decouple the spatial features into the respective regional and affine components in order to give a more detailed analysis of the regional motion differences

and to extract the dominant affine components. Since this decomposition cannot be validated, it is not addressed in this work. More detailed constructions are also possible to separate the space parameters, such as using separate tensors for each region to provide another type of regional analysis.

E. Understanding Pathology-specific Motion Abnormalities

Identification of cardiac motion abnormalities and classification of normal vs. abnormal motion patterns is an important step towards better understanding cardiac diseases that affect function. Given the formulation of the proposed methods, in which the motion is defined with a low number of parameters (the Polyaffine parameters), and given that the parameters can be aligned to a chosen reference space (using the methods described in Sec. II-B), different subjects can be compared, and models representing the mean motion of a population can be constructed (as in Sec. II-C) in order to analyse population-specific motion patterns. Mean motion models computed on different populations could be compared to the Polyaffine parameters from a patient-specific simulation to help to classify patients according to different diseases or different stages of a given disease, to aid in cases where there is uncertainty in the diagnosis of a specific patient. This classification could be combined with standard clinical measures to give a more informative picture of the patients condition. Also, by applying the mean model to a new subject (following the method described in Sec. II-C2), synthetic sequences can be generated to, for example, create a database of sequences with known ground-truth deformation to test cardiac models. A key advantage of the proposed mean motion model formulation is the ability to create full 3D sequences from the different models, to better understand the motion abnormalities. Such models could be used to complement existing methods to identify abnormal regions (strain or wall motion analysis for example), by providing more detail on what may be causing abnormal motion patterns in a population by visually analysing the motion dynamics.

III. VALIDATION

In order to test the methods described in the previous sections, 4D sequences of two populations were studied. The first; a data-set of healthy controls, to establish normal motion patterns, and the second; a data-set of patients with Tetralogy of Fallot, to establish motion patterns related to this pathology.

a) Healthy Volunteers: We illustrate these tools on 15 healthy adults (3 female, mean age \pm SD = 28 ± 5) obtained from the STACOM 2011 MICCAI cardiac motion tracking challenge database [34]. Steady-state free precession magnetic resonance images were acquired using a 3T scanner (Philips Achieva System, Philips Healthcare) in the short axis view covering entirely both ventricles (12-16 slices; isotropic in-plane resolution: $1.15 \times 1.15 \text{ mm}^2$ to $1.25 \times 1.25 \text{ mm}^2$; slice thickness: 8 mm ; 30 frames).

b) Tetralogy of Fallot Patients: To test the methods on pathological cases, we applied the described method to ten patients with repaired Tetralogy of Fallot (5 female, mean age \pm SD = 21 ± 7). These patients all had a full repair

early in infancy, resulting in the destruction of the pulmonary valves. Steady-state free precession magnetic resonance images were acquired for each patient in the short axis view covering entirely both ventricles (12-15 slices; isotropic in-plane resolution: $1.21 \times 1.21 \text{ mm}^2$ to $1.36 \times 1.36 \text{ mm}^2$; slice thickness: 8 mm ; 15-19 frames).

A. Validation of the Motion Tracking

Since the proposed population-based motion analysis is based on the initial motion tracking, the accuracy of the motion tracking has a large impact on the subsequent analysis. In order to examine the accuracy of the motion tracking, the cardiac-specific PolyLogDemons method was applied to the STACOM 2011 MICCAI cardiac motion tracking challenge database [34]. The primary goal here was to validate the choice of parameters used in the model to determine if the input parameters, the number of regions, and the choice of weight functions, provide sufficiently accurate tracking on a level equivalent to state-of-the-art methods.

The STACOM 2011 cardiac motion tracking challenge was open to all willing participants. Of the four competitors that applied for the challenge, only two applied their methods to the cine MR sequences. These were the Temporal Diffeomorphic Free Form Deformation (TDFFD) method of De Craene et. al [8] and the iLogDemons method of Mansi et. al [27], [4]. The principle behind the TDFFD method is to model the deformations as B-splines. The iLogDemons is rather a gradient-driven optical flow method based on the Demons method. As with the method used in this work, both of these methods require a mask of the myocardium as input.

Evaluating the accuracy of the provided ground truth landmark tracking is not the objective of this work. Rather, we are interested in quantitatively comparing the registration results of the proposed method to state-of-the-art methods applied to the same sequences. Given that the ground truth computation is subject to errors from the landmark allocation itself (inter- and intra-observer bias), and in alignment between the tagged MR and cine MR sequences (subjects may move between acquisitions), and noting the large slice thickness (8 mm), we are interested in evaluating how greatly the errors vary from the other methods, rather than the absolute error itself. Furthermore, we expect higher errors than for ECHO or tagged MR for instance, since there is little or no texture information in the myocardium to drive the registration [35].

We applied the same work-flow for computing the error as described in [34] to be able to directly compare the errors between the different methods. Based on the experiments described in [14], [13], the chosen parameters were $\alpha = 1$, $\beta = 1$, $\kappa = 0.5$, $\mu = 1$, with PSS weights, and three multi-resolution levels. The l^2 norm error, between the annotated landmarks at end diastole and end systole and the deformed landmarks at the corresponding frames, was computed. The median error value of the two time points was 3.07 for the proposed method. This can be directly compared to the median values for the iLogDemons algorithm (4.82) and for the TDFFD algorithm (3.17). Therefore, the errors are on the same order of magnitude as these methods. The errors are

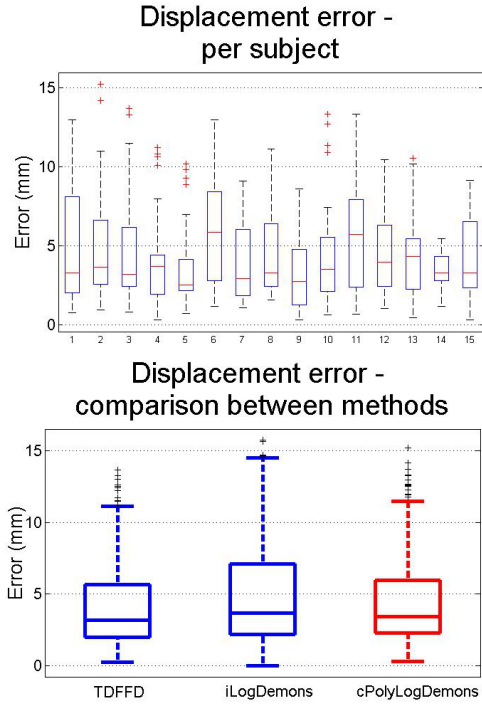


Fig. 8. Box plot of errors per subject from the 12 4D annotated landmarks at end diastole and end systole (top). The average over the population is plotted for the TDFFD algorithm, iLogDemons algorithm, and the cardiac PolyLogDemons method (bottom). The cardiac PolyLogDemons method is on the same order of accuracy as both the TDFFD and the iLogDemons method.

plotted in Fig. 8 (top) for each of the 15 volunteers. The error over all the subjects is shown in Fig. 8 (bottom) along with the box plots for the TDFFD and iLogDemons methods for comparison.

B. Validation of the Mean Motion Model

A mean normal (healthy) motion model was computed from the 15 healthy volunteers. The first volunteer was chosen to be the reference, and all other subjects were aligned to this space. Note that the choice of reference is arbitrary since the parameter space is the space of affine transformations described in Cartesian coordinates. To visualise the mean motion models, the models were transported to the space of one of the volunteers. The first image was deformed by each motion model using the method described in Sec. II-C2. In order to analyse the results, we compare the mean models to both the Polyaffine model and to the original sequence. The images at three frames (frame 5, 10 and 20) are shown in Fig. 9. The healthy mean motion model gives a reasonable motion pattern, and actually gives smoother motion than the Polyaffine model. As shown in the second column for the Polyaffine sequence, the radial contraction is underestimated in the regions close to the junctions of the right ventricle with the left ventricle (as shown by the yellow arrow). The Tetralogy of Fallot mean motion model shows abnormal dynamics, with abnormal flattening of the left ventricle free wall towards the septum. This abnormal flattening at the septum has been observed in many Tetralogy of Fallot patients, and may be due to a high

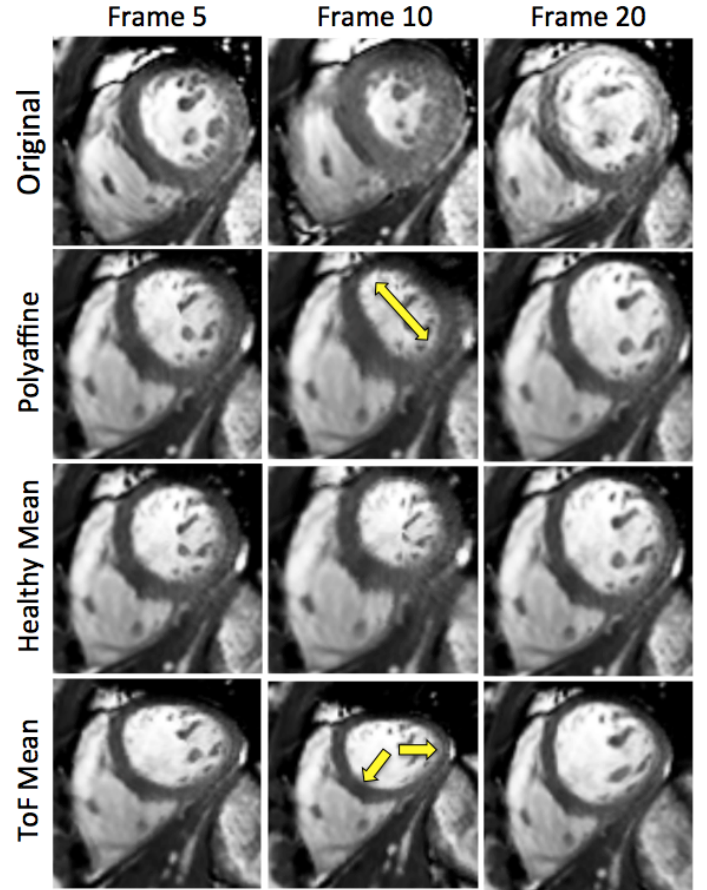


Fig. 9. Original sequence (top row), Polyaffine Model (second row), healthy mean motion model (third row) and Tetralogy of Fallot mean motion model (bottom row) at frame 5 (left column), 10 (centre) and 20 (right column) for one subject. The Polyaffine model underestimates the radial motion in some regions (indicated with the yellow arrow on frame 10), whereas the healthy mean motion captures smoother, more physiological motion. The Tetralogy of Fallot mean motion model shows abnormal motion dynamics, with inward motion at the left ventricle free wall and expansion, rather than contraction at the junctions of the left and right ventricles (shown in the bottom row, centre column with yellow arrows).

prevalence of post-repair right bundle branch block in the ECG (through a patch to repair the ventricular septal defect (VSD)) [36], the effect of the VSD-patch itself [37], and mediated through the persisting right ventricular volume and/or pressure load resulting from chronic pulmonary regurgitation/stenosis [38]. As a result, inter- and intra-ventricular dyssynchrony is highly prevalent and is regarded as an important contributor to both right ventricle and left ventricle dysfunction. More circumferential motion is also observed in the Tetralogy of Fallot mean motion sequence, as can be seen in the video of all of the sequences (see the supplementary video). Note that the key region of interest in these sequences is the myocardium of the left ventricle. Since the first image is deformed by each model, there are some motion artefacts in the blood pool caused by the trabeculae present in the first image. Also, the surrounding regions move due to the nature of the smooth Polyaffine weights.

A leave-one-out experiment was performed to test the ability of the mean model to approximate the motion of a new subject

drawn from the same population by using 14 subjects as the training set, and testing on the 15th subject. The residual error between the mean motion model (in the k^{th} subject space) and the original Polyaffine transformation parameters (before realignment) was computed for subjects in the training set as well as for subjects in the testing set. The errors are plotted over time in Fig. 10 for the training set (top) and the testing set (bottom). The errors are within the same range for the training and testing set, indicating that the model is able to predict the motion of healthy subjects.

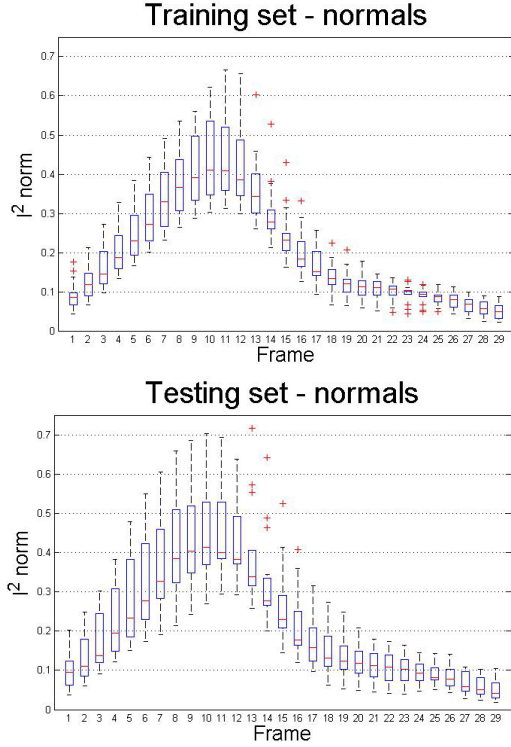


Fig. 10. Box plots of the residual norm between the Polyaffine transformation and the normal mean model transformation over time for subjects used in the training set (top), and for subjects not used in the training set (bottom) for the healthy volunteers.

Performing the same leave-one-out experiment on the Tetralogy of Fallot patients, we find that the residual errors are also not higher for the training set than for the testing set, as shown in Fig. 11. Therefore, this suggests that the mean Tetralogy of Fallot motion model is able to capture the motion dynamics of these patients reasonably well.

C. Analysis of the Predictive Power of the Mean Models

To test the ability of the mean normal model to capture the motion of Tetralogy of Fallot patients, the mean motion model was applied to each patient, and the weighted residual error between the model and the Polyaffine parameters was computed. Note that this distance measure was chosen to take into account the standard deviation of the data without requiring the computation of the inverse of the covariance matrix, as in the standard Mahalanobis distance (since the covariance matrix is ill-conditioned). The box plot of the error is given in Fig. 12 (top) alongside the box plots of the

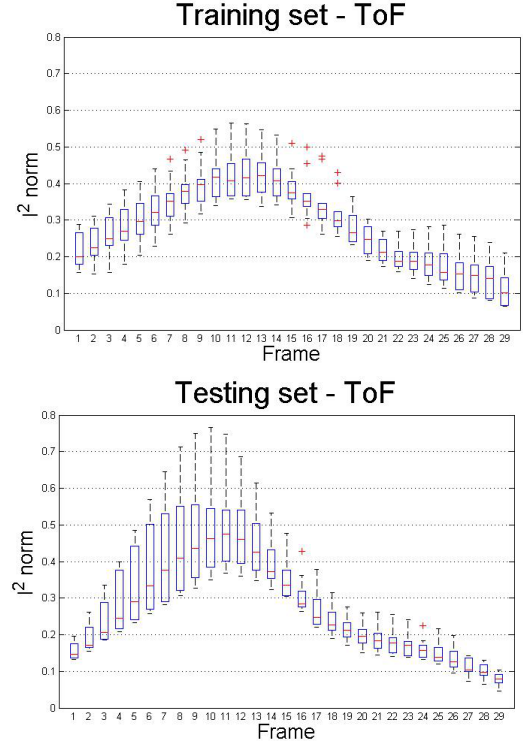


Fig. 11. Box plots of the residual norm between the Polyaffine transformation and the normal mean model transformation over time for subjects used in the training set (top), and for subjects not used in the training set (bottom) for the Tetralogy of Fallot patients.

training and testing set of the healthy volunteers. As expected, the errors are higher for the Tetralogy of Fallot patients, indicating that the healthy model is insufficient in capturing the motion dynamics in these patients. This suggests that there are abnormal motion dynamics observed in these patients that are not captured in the healthy model.

As a further test to quantify the differences between the populations and to overcome the issue of inverting the covariance matrix in the standard Mahalanobis distance computation, a projected Mahalanobis distance was computed. By projecting the parameters to a 1D plane, the Mahalanobis distance was computed directly from the projections γ_i given by:

$$\gamma_i = \frac{(V_i - \bar{V})^T (\bar{V}_{ToF} - \bar{V})}{\|\bar{V}_{ToF} - \bar{V}\|^2}, \quad (9)$$

where V_i is the vector of parameters for subject i ($i = 1, \dots, 25$), \bar{V} is the mean of the health controls, and \bar{V}_{ToF} is the mean of the Tetralogy of Fallot patients. In this formulation, if a subject is close to the mean of the healthy controls, then $\gamma \simeq 0$, if a subject is close to the mean of the Tetralogy of Fallot patients, then $\gamma \simeq 1$. A leave-one-out experiment was performed with the mean computed on 14 subjects and the projected Mahalanobis distance computed on the 15th subject, as well as on the Tetralogy of Fallot patients (see Fig. 12 - bottom). The mean for the healthy subjects was 0.8986, which can be compared to the Tetralogy of Fallot patients with 4.4513. Therefore, there is a statistically significant difference between the populations ($p < 0.05$).

In order to determine the classification power of the models

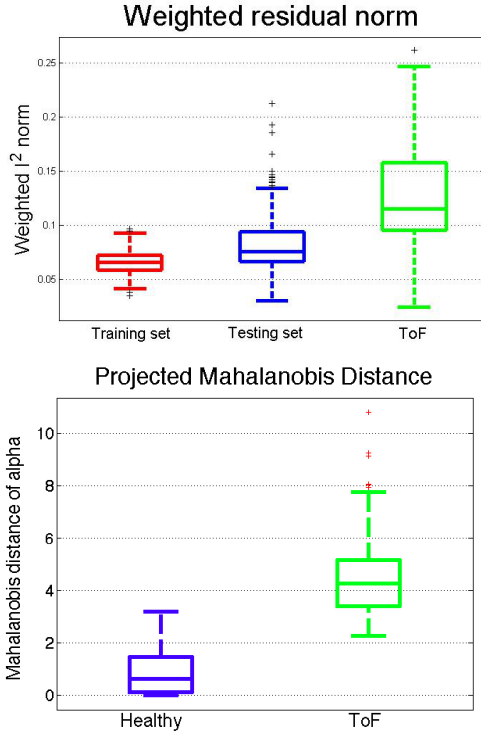


Fig. 12. Top: Box plots of the weighted residual norm (l^2 norm weighted with respect to the standard deviation of the mean model) between the Polyaffine transformation and the normal mean model transformation for subjects used in the training set (left - red), for healthy subjects not used in the training set (centre - blue) and for Tetralogy of Fallot patients (right - green). Bottom: Projected Mahalanobis distance measure on the alpha indices from Eq. 9 for the healthy subjects (blue) and Tetralogy of Fallot patients (green).

for each population, two-dimensional linear discriminant analysis (2DLDA) [39] was performed using the openly available Matlab code from Mathworks file exchange. Following the application of this method described in [15] to discriminate between different populations by setting the two dimensions in the 2DLDA formulation as the spatial and temporal dimensions, a classifier was trained in a leave-one-out nature on all the transformations. The subjects were labelled as either healthy or Tetralogy of Fallot, and the classifier was tested on the excluded subject. Setting the parameters p and q for the dimensions of the reduced subspace as $p = q = 15$, and running a k-Nearest-Neighbours classifier on the testing subject (with 3 neighbours), the classifier correctly predicted the population group with 100% accuracy. Reducing the dimension of the acquired subspace to $p = q = 5$ yielded a classification accuracy of 92%.

D. Validation of the Tensor Decomposition

In order to capture the spatial and temporal aspects of the motion, a 3-way tensor of all the spatial parameters (affine parameters for all regions) stacked along time and for all subjects was constructed. Stacking the tensor in this way gives a decomposition in terms of the spatial components and temporal components separately (though other constructions are also possible). The Tucker tensor decomposition was performed using the Sandia National Laboratories Tensor Toolbox [40],

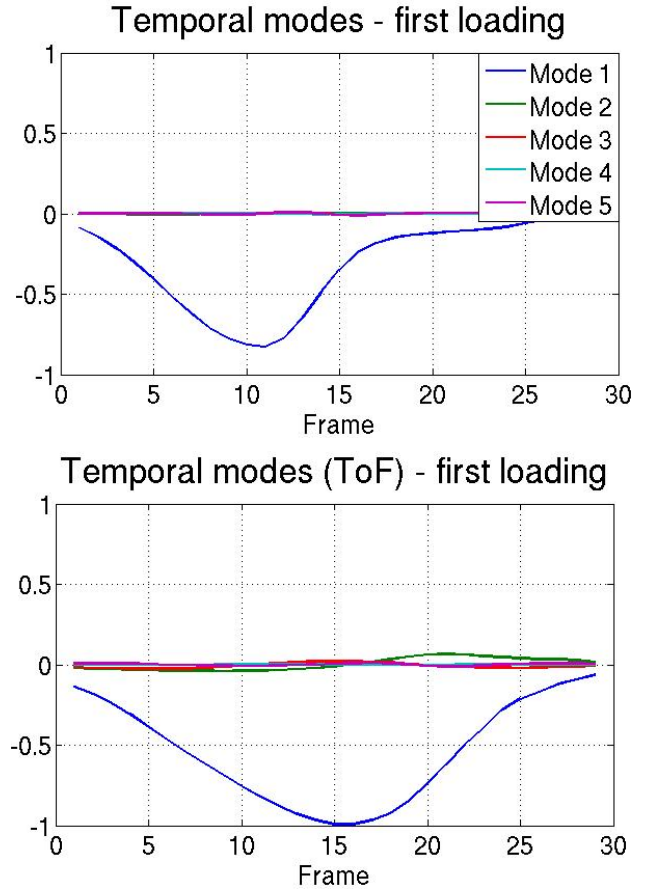


Fig. 14. Temporal modes for the first five loadings corresponding to the first spatial mode for the healthy controls (top), and for the Tetralogy of Fallot patients (bottom). The first loadings show interesting differences for the Tetralogy of Fallot patients, with slower contraction and slower filling than for the healthy subjects.

implemented in MATLAB. Using these tools, the three-way tensors of $[space \times time \times subject]$ of the healthy controls and of the Tetralogy of Fallot patients, were decomposed into factor and core matrices. Using the Tucker decomposition, we obtain temporal bases for each loading related to the spatial bases.

The first two spatial modes for the healthy controls are shown in Fig. 13 (left) and for the Tetralogy of Fallot patients (right). The first mode for the healthy controls appears to capture both the radial contraction and the circumferential motion (shown in block yellow arrows). The Tetralogy of Fallot modes, on the other hand, appear to capture a translation of the free wall and septal wall towards the right ventricle (RV). This abnormal motion is evident in the image sequences of these patients.

The temporal modes are shown in Fig. 14 (corresponding to the first loading of the spatial modes) for the first five temporal modes, shown for the healthy controls (top) and Tetralogy of Fallot patients (bottom). These plots show the dominant temporal mode in blue, which corresponds well to the expected volume curve for healthy subjects, but may indicate slower contraction and slower filling for the Tetralogy of Fallot patients.

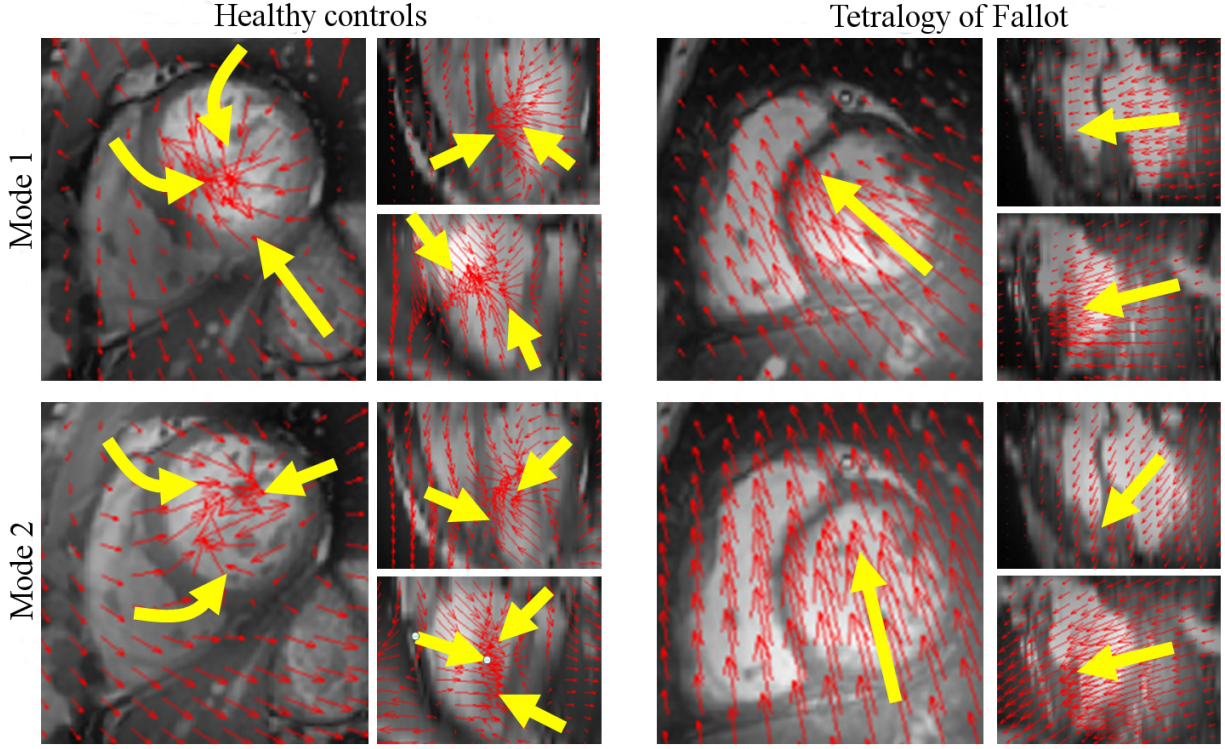


Fig. 13. Three views of the first (top row) and second (bottom row) spatial modes for the healthy controls (left) and for the Tetralogy of Fallot patients (right). The modes for the healthy controls represent the radial contraction and circumferential motion, whereas the modes for the Tetralogy of Fallot patients represent the translation towards the right ventricle. Yellow arrows indicate the general direction of motion.

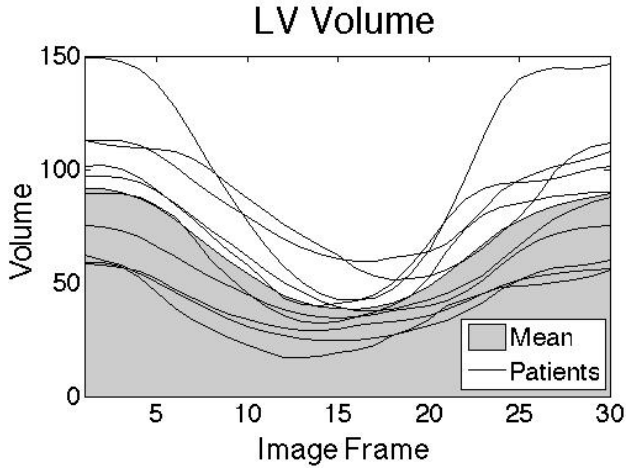


Fig. 15. Left ventricle volume curves over time for the 10 Tetralogy of Fallot patients (in black) and the mean volume (shaded gray), which shows similar trends to the principal temporal mode shown in Fig. 14.

The volume curves for the Tetralogy of Fallot patients were computed for the left ventricle from 4D segmentation performed using the MedViso Segment software (openly available for research purposes [41]), as shown in Fig. 15. The mean curve (shown in gray) corresponds well to the dominant temporal mode for these patients shown in Fig. 14, which indicates that the end systolic phase is delayed compared to healthy subjects and confirms the findings for the first temporal modes.

E. Implementation and Processing Time

The implementation of the described methods requires a number of key steps, including pre-processing of the images, segmentation of the ventricle and division into the Polyaffine regions (the AHA regions), computation of the Polyaffine weights, execution of the motion tracking registration, tensor construction, and model reduction. The pre-processing of the images requires image contrast enhancement and isotropic voxel resampling for the LogDemons part of the registration (which is computed on a regular grid and thus requires the voxel size to be isotropic). The left ventricle needs to be segmented in order to define the AHA regions and to compute the Polyaffine weights (which can both be performed automatically given the segmentation as well as the labelling of the apex, base and the junctions between the left and right ventricles). Given the regions and the weights, the registration algorithm can be launched by sequentially computing the registration of the current frame to the reference frame (the end-diastolic frame) and initialising the registration with the previous frame-to-reference velocity field. The registration is computed using ITK and C++ functions, building from the ITK implementation of the LogDemons algorithm [42]. Constructing the tensor from the computed parameters is simply tensor manipulation and stacking. The model reduction on the constructed tensor was performed in Matlab using the tensor decomposition methods from Sandia National Laboratories [33]. In terms of implementation, the key bottleneck is in the segmentation of the left ventricle, which in this case

was performed using semi-automatic tools which requires on average 30 minutes per surface, and can take more than an hour in difficult cases (where there is large slice misalignment for example). Computation of the regions and weights is fully automatic and reasonably fast with C++ programming. The registration itself takes around 30 minutes per registration, therefore, given the sequential nature of the computation for a full heart cycle, it can take several hours for one cardiac sequence. Finally, the tensor construction and decomposition can be quickly computed in Matlab (in minutes). Ideally, the steps could be automated and the computation time could be improved by taking advantage of parallel computing architectures and by minimising the need for user intervention with fully automatic segmentation and landmark detection tools.

IV. DISCUSSION AND PERSPECTIVES

In the current work, we proposed to use statistical methods to create a mean motion model, and to extract the modes of variation in terms of the spatial and temporal components using image registration, statistical model reduction, and tensor decomposition methods. The proposed method uses the parameters of the cardiac PolyLogDemons algorithm [14], which gives transformations described by a small number of parameters extracted from cine MR image sequences. The analysis of cine MR images was chosen specifically for the frequency of use in standard clinical practice (no radiation risks and short axis cine loops are standard in the clinical protocol for many diseases), which allows for more widespread application of the method, as well as for the good spatial resolution, which allows for better visualisation of the myocardial borders.

The proposed methods were applied to motion tracking of the left ventricle only. The reasons behind this choice are described in Sec. IV-1. Spatio-temporal alignment of the parameters was performed prior to the analysis. The techniques used for the alignment are further discussed in Sec. IV-2. A mean motion model was constructed by taking the average of the aligned parameters. The mean model construction is further discussed in Sec. IV-3. From the obtained Polyaffine transformation parameters, statistical model reduction was applied to the parameters after spatio-temporal alignment. The statistical methods used are discussed in Sec. IV-4. A brief discussion on an interpretation of the results and how the proposed methods can be used to gain further insight into pathology-specific functional abnormalities is described in Sec. IV-5.

1) *Bi-ventricle Motion Analysis*: The present work is focused primarily on the left ventricle since the current ground truth data for the motion is available for the left ventricle only. As a proof of concept article it does not seem viable to draw any conclusions from the statistical model (namely the tensor decomposition) built from the motion tracking (the cardiac PolyLogDemons method) without having some validation of the motion tracking parameters. Furthermore, investigating left ventricular dynamics in Tetralogy of Fallot subjects is of interest for clinicians since left ventricle dysfunction has been identified as a predictor for mortality in adults with this disease

[43]. For future applications to clinical data the method can be extended to bi-ventricular (i.e. left and right ventricle) motion analysis, provided that anatomically relevant weights can be defined on the right ventricle.

2) *Parameter Realignment*: Spatio-temporal alignment was performed by making use of the trace of the affine matrices to align the transformations temporally, and making use of the prolate spheroidal coordinates to align the transformations spatially, in an anatomically meaningful manner. The temporal realignment used in this work was a simple rigid transformation of the parameters, to translate the parameters regionally to the mean peak of the trace. However, this may have the effect of removing relevant parameters. Therefore, a non-rigid transformation may allow more intuitive analyses by fixing the first frame and stretching (or shrinking) the parameters to the mean peak of the trace, and stretching (or shrinking) to the relaxation phase, similar to the temporal alignment described in [44]. Both the temporal and spatial alignment were performed at a regional level. This was due to the fact that we are interested in region-wise analyses, rather than global analyses. However, it is also possible to temporally align the transformations to a global peak frame. Spatially, however, the current formulation relies on the regional construction. The parameters could be spatially aligned by the barycentre of the ventricle, however, further data-points would be needed such as the base, apex and the junctions between the left and right ventricles.

3) *Mean Motion Model Construction*: From the mean motion model constructed from healthy subjects, we found that the model is able to approximate the motion of other healthy subjects reasonably well, however, the motion of Tetralogy of Fallot patients was not fully captured from this model. This was expected given the abnormal motion dynamics observed in Tetralogy of Fallot patients. Interestingly, the motion model constructed from 9 Tetralogy of Fallot patients was sufficient to capture the motion dynamics of the 10th subject (based on the results of the leave-one-out experiment). Given that these patients have followed different treatment paths, and are at different stages due to the range of patient age (some patients are children, others adults), we expected to require a larger number of patients to capture the dynamics in this population. However, the related factor between these patients is a missing pulmonary valve, caused by the initial surgery. Therefore, the regurgitated blood from the pulmonary artery may be the key factor affecting the motion dynamics of these patients. Ideally, given more patients with similar treatment history, treatment-specific models could be constructed to distinguish the key differences in the motion dynamics for patients treated with one treatment over another. In future work, the mean motion models could also be used to generate synthetic sequences for new subjects in which only the end diastolic image is available.

4) *Statistical Model Reduction*: A number of statistical model reduction techniques were used in this work to attempt to understand the underlying motion variability of the two populations. We extend on the work in [23], in which model reduction is performed block-wise on matrices stacked either column- or row-wise to obtain temporal and spatial population bases. In the current work, we perform model reduction on the

full tensor using the Tucker decomposition to obtain spatial and temporal bases directly. The advantage of performing model reduction directly on the tensor is that the bases are decoupled and can be recombined independently. However, tensor decomposition is not straightforward and requires numerical methods for estimating the modes, compared to matrix decomposition which can be performed simply with SVD, PCA, partial least squares, etc. The tensor decomposition used in this work was the Tucker decomposition, though a number of other methods for decomposing a tensor have been proposed and could be applied (such as CP decomposition, 2DPCA, etc.).

5) *Population-wide Analysis of 4D Motion*: The validation section (Sec. III) showed the ability of the proposed population-based analysis methods for extracting the dominant patterns observed in a population. For healthy subjects, these were the radial, circumferential, and longitudinal motion representing the key spatial patterns, and the temporal trigger which follows the pattern of the volume curve in the ventricle. In contrast, the typical motion patterns for Tetralogy of Fallot patients are less well known. The presented method identified the dominant spatial pattern in these hearts as being a translation from the left ventricle free wall to the septum, which is consistent with qualitative analysis of the motion visible for some patients, in which the left ventricle appears to be dragged towards the right ventricle. The identified temporal trigger indicates that there is a slower contraction of the left ventricle, which was verified by computing the volume over time for all patients using image segmentation. These experiments may provide further insight into the function of the left ventricle for Tetralogy of Fallot patients, with more efficiency than manual analysis (i.e. 4D segmentation of several hearts may be time-consuming and subject to user bias) and may provide more detail than simpler methods (such as 1D measures of strain or wall thickening).

6) *4-Way Tensor Decomposition*: The goal of the proposed methodology is to gain insight into pathology-specific functional behaviour by analysing the motion to extract information that may currently be unknown. The experiments described in the validation section showed that both the mean motion model and the tensor decomposition of grouped motion parameters give meaningful results in terms of extracting already known information automatically from the data (such as spatial and temporal behaviour). In future work, the analysis will go deeper to investigate and compare different populations. In order to analyse the spatial motion features independently (in terms of regional and affine components), 4-way Tucker Decomposition can be performed on tensors stacked by [affine parameters \times region \times time \times subject]. Further work is required for such analysis in terms of optimising for the number of dimensions needed in the reduced model and describing the parameters in a more physiological sense.

V. CONCLUSION

In this work, we propose a novel method for analysing population-wide motion patterns in terms of the dominant spatial and temporal patterns. Using a Polyaffine model to

describe the transformation from one image to another, we perform group-wise analysis on these parameters to construct a mean motion model by averaging spatio-temporally aligned transformation parameters from a group of subjects in the same population (healthy or pathological). Further population-based statistical analyses were performed by computing a tensor decomposition on the transformation parameters directly. From the analyses we obtained meaningful differences between the healthy population and the pathological population of Tetralogy of Fallot patients in terms of motion patterns (both spatially and temporally) and were able to detect potentially meaningful motion information for Tetralogy of Fallot patients such as an abnormal triggering of contraction.

APPENDIX

A. Tucker Tensor Decomposition using Alternating Least Squares

The alternating least squares (ALS) method for computing a rank- (R_1, \dots, R_N) Tucker decomposition is described briefly here. Note that the n -Rank of a tensor \mathcal{T} is not the same as the rank of a tensor (the minimum number of rank-one components). R_n is the column rank of the n^{th} axis of \mathcal{T} . In this work, we take $R_n = \text{rank}_n(\mathcal{T})$ to find an exact Tucker decomposition of \mathcal{T} , rather than computing a truncated Tucker decomposition (which gives an inexact decomposition rather than the truncation of an exact decomposition).

The ALS method essentially approximates the factors that minimise the least squares criterion:

$$C(\mathcal{T}) =: \min_{\mathcal{G}, \mathcal{A}^{(n)}} \|\mathcal{T} - [[\mathcal{G}; \mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots, \mathcal{A}^{(N)}]]\| \quad (10)$$

given $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times \dots \times R_N}$ and $\mathcal{A}^{(n)} \in \mathbb{R}^{I_n \times R_n}$ is column-wise orthogonal for all n . The algorithm for computing a rank- (R_1, \dots, R_N) decomposition is given is Algorithm 1 [33].

Algorithm 1 ALS rank- (R_1, R_2, \dots, R_N) Tucker decomposition

Require: Initialisation: $\mathcal{A}^{(n)} \in \mathbb{R}^{I_n \times R}$ for $n = 1, 2, \dots, N$
loop {until convergence}
 for $n = 1, \dots, N$ **do**
 $y \leftarrow \mathcal{T} \times_1 \mathcal{A}^{(1)T} \dots \times_{n-1} \mathcal{A}^{(n-1)T} \times_{n+1} \mathcal{A}^{(n+1)T} \dots \times_N \mathcal{A}^{(N)T}$
 $\mathcal{A}^{(n)} \leftarrow$ the $R^{(n)}$ leading singular vectors of y
 $\mathcal{G} \leftarrow \mathcal{T} \times_1 \mathcal{A}^{(1)T} \times_2 \mathcal{A}^{(2)T} \dots \times_N \mathcal{A}^{(N)T}$
return $\mathcal{G}, \mathcal{A}^{(1)}, \dots, \mathcal{A}^{(N)}$

ACKNOWLEDGMENTS

This work was carried out predominantly at the INRIA research institute and was partially funded by the Care4Me ITEA2 project and the MD Paedegree European project ICT-2011.5.2 grant no. 600932. The authors would like to thank the organisers of the STACOM 2011 cardiac motion tracking challenge, Nicolas Toussaint and Christof Seiler for their help with the prolate spheroidal coordinates and the Polyaffine algorithm, and the Sandia National Laboratories for providing open source tools for tensor manipulation and decomposition.

The authors would also like to kindly thank the reviewers for their useful feedback and comments which helped to improve this work.

REFERENCES

- [1] A. F. Frangi, W. J. Niessen, and M. A. Viergever, "Three-dimensional modeling for functional analysis of cardiac images, a review," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 1, pp. 2–5, 2001.
- [2] H. Wang and A. A. Amini, "Cardiac motion and deformation recovery from MRI: a review," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 2, pp. 487–503, 2012.
- [3] L. Tautz, A. Hennemuth, and H.-O. Peitgen, "Motion analysis with quadrature filter based registration of tagged MRI sequences," in *Proc. STACOM MICCAI Workshop*, ser. LNCS. Springer, 2011.
- [4] T. Mansi, X. Pennec, M. Sermesant, H. Delingette, and N. Ayache, "iLogDemons: A demons-based registration algorithm for tracking incompressible elastic biological tissues," *Int J. of Comp Vision*, vol. 92, no. 1, 2011.
- [5] R. Chandrashekar, R. Mohiaddin, and D. Rueckert, "Cardiac motion tracking in tagged MR images using a 4D B-spline motion model and nonrigid image registration," in *IEEE Int. Symp. on Biomed. Imaging: Nano to Macro*, 2004, pp. 468–471.
- [6] W. Shi, X. Zhuang, H. Wang, S. Duckett, D. V. Luong, C. Tobon-Gomez, K. Tung, P. J. Edwards, K. S. Rhode, R. S. Razavi et al., "A comprehensive cardiac motion estimation framework using both untagged and 3D tagged MR images based on nonrigid registration," *IEEE Trans. Med. Imaging*, vol. 31, no. 6, pp. 1263–1275, 2012.
- [7] B. Heyde, D. Barbosa, P. Claus, F. Maes, and J. D'hooge, "Three-dimensional cardiac motion estimation based on non-rigid image registration using a novel transformation model adapted to the heart," in *Proc. STACOM MICCAI Workshop*, ser. LNCS, 2012.
- [8] M. De-Craene, C. Tobon-Gomez, C. Butakoff, N. Duchateau, G. Piella, K. Rhode, and A. Frangi, "Temporal diffeomorphic free form deformation (TDDFD) applied to motion and deformation quantification of tagged MRI sequences," in *Proc. STACOM MICCAI Workshop*, ser. LNCS. Springer, 2011.
- [9] O. Commowick, V. Arsigny, A. Isambert, J. Costa, F. Dhermain, F. Bidault, P.-Y. Bondiau, N. Ayache, and G. Malandain, "An efficient locally affine framework for the smooth registration of anatomical structures," *Med. Image Anal.*, vol. 12, no. 4, 2008.
- [10] V. Arsigny, O. Commowick, N. Ayache, and X. Pennec, "A fast and log-euclidean polyaffine framework for locally linear registration," *Journal of Mathematical Imaging and Vision*, vol. 33, no. 2, pp. 222–238, 2009.
- [11] W. Zhang, J. Noble, and J. Brady, "Spatio-temporal registration of real time 3D ultrasound to cardiovascular MR sequences," *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2007*, pp. 343–350, 2007.
- [12] M. S. Hansen, S. S. Thorup, and S. K. Warfield, "Polyaffine parametrization of image registration based on geodesic flows," in *Proc. MMBIA Workshop*. IEEE, 2012, pp. 289–295.
- [13] K. McLeod, C. Seiler, M. Sermesant, and X. Pennec, "A near-incompressible poly-affine motion model for cardiac function analysis," in *Proc. STACOM MICCAI Workshop*, ser. LNCS. Springer, 2012.
- [14] K. McLeod, C. Seiler, N. Toussaint, M. Sermesant, and X. Pennec, "Regional analysis of left ventricle function using a cardiac-specific polyaffine motion model," in *Proc. FIMH'13*, 2013.
- [15] Z. Qian, Q. Liu, D. N. Metaxas, and L. Axel, "Identifying regional cardiac abnormalities from myocardial strains using nontracking-based strain estimation and spatio-temporal tensor analysis," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 12, pp. 2017–2029, 2011.
- [16] A. Rao, G. I. Sanchez-Ortiz, R. Chandrashekar, M. Lorenzo-Valdés, R. Mohiaddin, and D. Rueckert, "Comparison of cardiac motion across subjects using non-rigid registration," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2002*. Springer, 2002, pp. 722–729.
- [17] —, "Construction of a cardiac motion atlas from mr using non-rigid registration," in *Functional Imaging and Modeling of the Heart*. Springer, 2003, pp. 141–150.
- [18] D. Perperidis, M. Lorenzo-Valdés, R. Chandrashekar, A. Rao, R. Mohiaddin, G. Sanchez-Ortiz, and D. Rueckert, "Building a 4D atlas of the cardiac anatomy and motion using MR imaging," in *Biomedical Imaging: Nano to Macro*, 2004. IEEE International Symposium on. IEEE, 2004, pp. 412–415.
- [19] D. Perperidis, A. Rao, M. Lorenzo-Valdés, R. Mohiaddin, and D. Rueckert, "Spatio-temporal alignment of 4D cardiac MR images," in *Functional Imaging and Modeling of the Heart*. Springer, 2003, pp. 205–214.
- [20] N. F. Rougon, C. Petitjean, and F. J. Preteux, "Building and using a statistical 3D motion atlas for analyzing myocardial contraction in MRI," in *Medical Imaging 2004*. International Society for Optics and Photonics, 2004, pp. 253–264.
- [21] R. Chandrashekar, A. Rao, G. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert, "Construction of a statistical model for cardiac motion analysis using nonrigid image registration," in *Information Processing in Medical Imaging*. Springer, 2003, pp. 599–610.
- [22] Z. Xue, D. Shen, and C. Davatzikos, "Statistical representation of high-dimensional deformation fields with application to statistically constrained 3D warping," *Medical Image Analysis*, vol. 10, no. 5, pp. 740–751, 2006.
- [23] K. McLeod, C. Seiler, M. Sermesant, and X. Pennec, "Spatio-temporal dimension reduction of cardiac motion for group-wise analysis and statistical testing," in *MICCAI - Medical Image Computing and Computer Assisted Intervention - 2013*, ser. Lecture Notes in Computer Science. Springer, Heidelberg, 2013.
- [24] V. Arsigny, O. Commowick, N. Ayache, and X. Pennec, "A fast and log-euclidean polyaffine framework for locally linear registration," *J. of Math. Imaging and Vision*, vol. 33, no. 2, 2009.
- [25] C. Seiler, X. Pennec, and M. Reyes, "Capturing the multiscale anatomical shape variability with polyaffine transformation trees," *Med. Image Anal.*, 2012.
- [26] N. Toussaint, C. T. Stoeck, T. Schaeffter, S. Kozerke, M. Sermesant, and P. G. Batchelor, "In vivo human cardiac fibre architecture estimation using shape-based diffusion tensor processing," *Medical Image Analysis*, pp. –, 2013.
- [27] K. McLeod, A. Prakosa, T. Mansi, M. Sermesant, and X. Pennec, "An incompressible log-domain demons algorithm for tracking heart tissue," in *Proc. STACOM MICCAI Workshop*, ser. LNCS. Springer, 2011.
- [28] L. Le Folgoc, H. Delingette, A. Criminisi, and N. Ayache, "Sparse bayesian registration," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2014*. Springer, 2014.
- [29] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of eckart-young decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.
- [30] R. A. Harshman, "Foundations of the parafac procedure: models and conditions for an" explanatory" multimodal factor analysis," 1970.
- [31] L. R. Tucker, "Implications of factor analysis of three-way matrices for measurement of change," *Problems in measuring change*, pp. 122–137, 1963.
- [32] —, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.
- [33] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [34] C. Tobon-Gomez, M. De-Craene, K. McLeod, T. Lautz, W. Shi, A. Hennemuth, A. Prakosa, H. Wang, G. Carr-White, S. Kapetanakis, A. Lutz, V. Rasche, T. Schaeffter, C. Butakoff, O. Friman, T. Mansi, M. Sermesant, X. Zhuang, S. Ourselin, H.-O. Peitgen, X. Pennec, R. Razavi, D. Reuckert, A. Frangi, and K. Rhode, "Benchmarking framework for myocardial tracking and deformation algorithms: An open access database," *Med. Image Anal.*, 2013.
- [35] H. Wang and A. Amini, "Cardiac motion and deformation recovery from MRI: a review," *IEEE Trans. Med. Imaging*, vol. 31, no. 2, 2012.
- [36] L. Jing, C. M. Haggerty, J. D. Suever, S. Alhadad, A. Prakash, F. Cecchin, O. Skrinjar, T. Geva, A. J. Powell, and B. K. Fornwalt, "Patients with repaired tetralogy of fallot suffer from intra- and inter-ventricular cardiac dyssynchrony: a cardiac magnetic resonance study," *European Heart Journal-Cardiovascular Imaging*, vol. 15, no. 12, pp. 1333–1343, 2014.
- [37] A. L. Knauth, K. Gauvreau, A. J. Powell, M. J. Landzberg, E. P. Walsh, J. E. Lock, P. J. del Nido, and T. Geva, "Ventricular size and function assessed by cardiac mri predict major adverse clinical outcomes late after tetralogy of fallot repair," *Heart*, vol. 94, no. 2, pp. 211–216, 2008.
- [38] W. Hui, C. Slorach, A. Dragulescu, L. Mertens, B. Bijnsens, and M. K. Friedberg, "Mechanisms of right ventricular electromechanical dyssynchrony and mechanical inefficiency in children after repair of tetralogy of fallot," *Circulation: Cardiovascular Imaging*, pp. CIRCIMAGING–113, 2014.
- [39] Z. Liang, Y. Li, and P. Shi, "A note on two-dimensional linear discriminant analysis," *Pattern Recognition Letters*, vol. 29, no. 16, pp. 2122–2128, 2008.

- [40] B. W. Bader, T. G. Kolda et al., "Matlab tensor toolbox version 2.5," Available online, January 2012. [Online]. Available: <http://www.sandia.gov/~tgkolda/TensorToolbox/>
- [41] E. Heiberg, J. Sjögren, M. Ugander, M. Carlsson, H. Engblom, and H. Arheden, "Design and validation of segment-freely available software for cardiovascular image analysis," BMC medical imaging, vol. 10, no. 1, p. 1, 2010.
- [42] F. Dru, T. Vercauteren et al., "An itk implementation of the symmetric log-domain diffeomorphic demons algorithm," 2009.
- [43] A. Ghai, C. Silversides, L. Harris, G. D. Webb, S. C. Siu, and J. Therrien, "Left ventricular dysfunction is a risk factor for sudden cardiac death in adults late after repair of tetralogy of fallot," Journal of the American College of Cardiology, vol. 40, no. 9, pp. 1675–1680, 2002.
- [44] D. Perperidis, R. H. Mohiaddin, and D. Rueckert, "Spatio-temporal free-form registration of cardiac mr image sequences," Medical image analysis, vol. 9, no. 5, pp. 441–456, 2005.